Title page

Obesity is a risk factor for acute promyelocytic leukemia: evidence from population and cross-sectional studies and correlation with flt3 mutations and polyunsaturated fatty acid metabolism

Luca Mazzarella*[1,2], Edoardo Botteri[3], Anthony Matthews[4], Elena Gatti[1], Davide Di Salvatore[5], Vincenzo Bagnardi[6], Massimo Breccia[7], Pau Montesinos[8], Teresa Bernal[9], Cristina Gil[10], Timothy J Ley[11] , Miguel Sanz[8], Krishnan Bhaskaran[4], Francesco Lo Coco[12], and Pier Giuseppe Pelicci*[1]

Affiliations

1 Dept. of Experimental Oncology, European Institute of Oncology, Milan, Italy; 2 Division of Experimental Therapeutics , European Institute of Oncology, Milan, Italy;
3 Cancer Registry of Norway, Oslo, Norway; 4 London School of Hygiene & Tropical Medicine, London, United Kingdom; 5 IFOM-FIRC institute of molecular oncology, Milan, Italy; 6 University of Milan-Bicocca, Milan, Italy; 7 Division of Hematology, Department of Cellular Biotechnologies and Hematology, Sapienza University, Rome, Italy; 8 Hematology Department, Hospital Universitario y Politécnico La Fe, Valencia, Spain; 9 Hospital Central de Asturias, Oviedo, Spain; 10 Hospital General, Alicante, Spain; 11 Department of Medicine, Division of Oncology, Washington University in St. Louis, Saint Louis, MO; 12 Department of Hematology, University Tor Vergata, Roma, Italy

Running head: obesity and risk of AML/APL, FLT3 muts and fat metabolism

*contact information for correspondence

luca.mazzarella@ieo.it and piergiuseppe.pelicci@ieo.it

Department of Experimental Oncology

Campus IFOM-IEO

Via Adamello 16 - 20139 Milan, Italy

Phone: +39 02 57489831

Fax: +39 02 94375990

Word count: 3218

Table count: 5

Figure count: 3

Supplementary files: 1

Acknowledgments

Funding: AIRC-Cariplo Foundation, Italian Ministry of Health, European Hematology Association, Wellcome Trust and Royal Society

This paper is dedicated to the memory of our wonderful colleague, Prof. Francesco Lo Coco, who recently passed away

Abstract

Obesity correlates with hematological malignances including leukemias, but risk of specific leukemia subtypes like Acute Promyelocytic Leukemia and underlying molecular mechanisms are poorly understood.

We explored multiple datasets for correlation between leukemia, Body Mass Index and molecular features. In a population-based study (n=5.2 million), we correlated Body Mass Index with promyelocytic, other acute myeloid, lymphoid or other leukemias. In cross-sectional studies, we tested body mass index deviation in promyelocytic leukemia trial cohorts from what expected based on national surveys. We interrogated The Cancer Genome Atlas for transcriptional signatures and mutations enriched in promyelocytic leukemia and/or obesity and confirmed correlation between body mass and FLT3 mutations in promyelocytic leukemia cohorts by logistic regression.

In the population-based study, Hazard Ratio *per* 5 kg/m^2 increase was: promyelocytic leukemia 1.44 (95% CI 1.0-2.08); non-promyelocytic acute myeloid leukemias 1.17 (1.10-1.26); lymphoid leukemias 1.04 (1.0-1.09); other 1.10 (1.04-1.15). In cross-sectional studies, body mass deviated significantly from expected (Italy $p<0.001$, Spain $p=0.011$, USA $p<0.001$). Promyelocytic leukemia showed upregulation of polyunsaturated fatty acid metabolism genes. Oddds of FLT3 mutations were higher in obese acute myeloid leukemias (Odds Ratio=2.4, $p=0.007$), whether promyelocytic or not, a correlation confirmed in the pooled promyelocytic leukemia cohorts (OR 1.22, 1.05-1.43 *per* 5 kg/m^2).

These results strengthen the evidence for obesity as a bona fide risk factor for myeloid leukemias and in particular APL. FLT3 mutations and polyunsaturated fatty acid metabolism may play a previously underappreciated role in obesity-associated leukemogenesis.

Introduction

The aetiology of Acute Myeloid Leukemia (AML) remains poorly understood. Genetic predisposition or clear exposure to environmental mutagenic agents (smoking, benzene, radiation, prior chemotherapy) can be demonstrated only in a minority of cases [1]. Age is an independent risk factor, probably linked to progressive mutation accumulation and clonal stem cell expansion accompanying aging [2]. Though obesity has recently emerged as a prominent risk factor for a variety of solid tumors[3], its impact on hematological neoplasms is comparatively less studied. A moderate but consistently positive correlation between body mass index (BMI) ad incidence of leukemias has been identified in observational studies[4–6]. Yet, none of the collected evidence has been considered sufficiently strong to consider obesity as a *bona fide* risk factor for AML [3,7]. Most studies did not distinguish between myeloid/lymphoid and acute/chronic forms, nor between genetic subtypes within each form. AML is recognized as a highly heterogeneous disease with genetically diverse subtypes [8]. Subtypes have radically different outcomes and, similarly, their risk may be differentially affected by environmental factors. Identification of subtype-specific risk associations, however, is made difficult by their rarity.

A genetic subset of AML, Acute Promyelocytic Leukemia (APL), is characterized by a specific chromomal translocation (t15;17), homogeneous biology and response to clinical agents All-Trans Retinoic Acid (ATRA) and Arsenic Trioxide, which have rendered it the most curable form of AML to date[9]. We previously demonstrated that the risk of relapse after ATRA/Idarubicin is significantly increased in overweight/obese APL patients[10]. In the present report, we investigated the association of overweight/obesity with risk of developing APL and other leukemias. We describe the results of multiple studies across four Western populations

with significantly different dietary regimens and prevalence of obesity. The studies were concordant in demonstrating increased risk of developing APL in overweight/obesity subjects. In an effort to generate mechanistic hypotheses to explain this relationship, we analyzed transcriptomic and mutational data from the AML project in The Cancer Genome Atlas (TCGA)[11] and identified alterations selectively associated with obesity and/or APL which may be involved in obesity-associated leukemogenesis.

Methods

Detailed methods are provided as supplementary material

UK population-based study: data collection and statistical methods

Methods for the UK population study were described in depth previously [6]. The study was approved by the London School of Hygiene and Tropical Medicine Ethics Committee. To identify outcomes of specific leukaemia sub-types, CPRD clinical records were searched for codes relating to specific leukemia subgroups. We controlled for multiple covariates at time of the BMI record(s): age, smoking status,; alcohol use, previous diabetes diagnosis, index of multiple deprivation, calendar period, and stratified by sex.. We excluded people with missing smoking (49 206/5.24 million [0.9%]) and alcohol status (394 196/5.24 million [7.5%]. All CIs are presented at the 95% level.

Cross-sectional studies: data collection and statistical methods

APL cases from Spain were extracted from the PETHEMA database to include 414 cases diagnosed between 1998 and 2012. APL cases from Italy were 134 adult patients treated with AIDA protocol included in the previously described cohort [10]. APL cases from USA included the entire cohort of the published AML TCGA project [11] (n=20) plus 22 additional APL cases, unselected for any clinical variable, diagnosed at Washington University (Expanded TCGA cohort). For all case cohorts, BMI was measured at the time of diagnosis.

Data collection was approved by the Research Ethics Board of each participating institution, as referenced [11–14]. Data source for expected BMI in the local population are described in the supplementary methods.

Expression data analysis

Expression data (RPKM matrix) were downloaded from the AML TCGA data portal. [11]. The Quantitative Set Analysis for Gene Expression method as implemented in the quSAGE package [15] in the R programming language (v 3.2.3) was used to conduct supervised gene set enrichment analysis. We focused on the KEGG collection as it is enriched for metabolism-associated gene annotations[16].

Mutational data analysis

For the analysis in the TCGA data, mutational data were retrieved from the TCGA AML paper[11] and AML driver genes (restricted to those with at least 2 mutations in the dataset) were downloaded from IntOgen [17].

For the analysis of the retrospective cohort, FLT3 Internal Tandem Duplication (ITD) mutational data were provided by the referring centers.

Role of the funding source

Funding sources had no role in study design, collection, analysis, and interpretation of data; report writing nor decision to submit the paper for publication

Results

Population-based cohort study in the UK

Overall characteristics of the 5.24 million UK adults included in this study have been described previously[6]. 5.833 subjects with a diagnosis of "leukemia" over the observational time were included in the present analysis. These events were further classified in the

following groups: "APL" (n=26), "non APL-AML" (n=1.012), lymphoid leukemias ("LL"; n=2.823) and "other" (n=1.972). Median time lapse between BMI measurement and diagnosis was similar across subgroups (APL: 1,810 days; AML: 2,280 days; LL: 1,928 days; other: 1,894 days).

We fit BMI as a three-knot cubic spline and as a linear term. There was no evidence of non-linearity (p=0.94), suggesting that the relationship was best described by the linear model. After adjusting for covariates, *per* each 5 kg/m^2 increase we obtained hazard ratios (HR) of 1.44 for APL (95% confidence interval (CI) 1.0-2.08), 1.17 for non APL-AML (95% CI 1.10-1.26), 1.04 for LL (95% CI 1.0-1.09) and 1.10 for other leukemias (95% CI 1.04-1.15) (Figure 1 and Table 1). Stratification by gender suggested a stronger effect for male gender in APL (HR 1.82, 95% CI 1.10- 3.00 vs female HR 1.19, 95%CI 0.67-1.98), although the sample size becomes very small (n=13 each). Together, these results suggest that higher BMI is associated with increased risk of all sub-types of leukemia, particularly APL.

Cross-sectional studies in Italian, Spanish and USA trial cohorts

Though APL showed the strongest association with higher BMI in the cohort analysis described above, results were not conclusive due to the small number of cases identified (n=26) and the consequently wide confidence intervals. To strengthen the evidence, we carried out retrospective case-control studies using cohorts of APL patients from national registries of clinical trials from Spain (PETHEMA) and Italy (GIMEMA) and patients from the USA-based AML genome sequencing study (the AML TCGA cohort with 22 additional cases characterized at Washington University-St Louis). In all three groups, APL diagnosis was established using gold standard diagnostic procedures.

Demographic characteristics of the three case cohorts (Italy n=134, Spain n=414 and USA n=42) are described in Table 2. Gender (female 53.0%, 55.2%, 50% respectively) and age (median of 45, 45 and 47 respectively) were similarly represented. Information on ethnicity

was unavailable for the Spanish and Italian cohorts, whereas white, black and hispanic ethnicities were represented in the USA cohort.

To generate control groups for comparison, we obtained anthropometric data from epidemiological surveys of the general population in the different countries. As the prevalence of obesity has increased dramatically in most countries in the last decades (especially USA), we obtained data that were as close as possible to the median year of diagnosis (2002 for Italy, 2003 for Spain, 2007 for USA) (Table 3 and Methods).

In all three cohorts, there was strong evidence that the observed BMI distribution for cases across WHO BMI classes was different from that expected under the null hypothesis of no association (Italy $p<0.001$, Spain $p=0.011$, USA $p<0.001$; Table 3) in gender-, age- and ethnicity- (for USA) matched controls. In particular, in all 3 datasets, there were more cases than expected in the higher BMI groups, irrespective of gender in all cohorts apart from Spain, in which significance was not reached for males ($p=0.130$), despite a similar trend (Table 3).

Correlation of TCGA transcriptomics data with BMI and AML subtype

The availability of the TCGA dataset prompted us to search for signatures that could suggest a mechanistic rationale for the association between APL and obesity. We interrogated available AML transcriptomes with supervised gene set enrichment analysis using quSage [15]. Focusing on the KEGG gene set collection, APL was associated with increased activity of 13 and decreased activity of 64 out of 186 gene sets (figure 2A, table 4 and supplementary table S1). Intriguingly, among significantly upregulated gene sets we found pathways associated with the metabolism of long-chain unsaturated fatty acids (linoleic and arachidonic), which are precursors of eicosanoids mediating inflammation-associated cancers[18]. Also noticeable was the APL-associated upregulation of Insulin and Insulin-like Growth Factor (IGF1) receptors, but not leptin receptor (figure 2B); insulin signaling-associated pathways were also

specifically upregulated in obese vs non-obese APL patients ("type II diabetes mellitus" and "insulin signaling", supplementary table S1).

No pathway was significantly enriched in obese vs non-obese patients among non-M3 cases.

Correlation of mutational data data with BMI

We then asked whether obesity is associated with specific driver mutations in AML in the TCGA cohort. Out of 23 established driver genes mutated at least twice in the cohort, mutations in FLT3 were positively associated with obesity (33/88 obese vs 22/110 non-obese, p=0.007, Odds Ratio=2.4, False Discovery Rate (FDR)= 0.16, figure 3 and supplementary table S3). When we analysed the two main classes of FLT3 mutations separately (Tyrosine Kinase Domain, TKD and Internal Tandem Duplication, ITD), the association held statistically significant for ITD (24/88 obese vs 14/110 non-obese, p=0.01, Odds Ratio=2.6) but not for TKD (9/88 obese vs 8/110 non-obese, p=0.6). In APL, where all FLT3 mutations were ITD, the correlation remained statistically significant (6/12 obese vs 0/8 non-obese, p=0.04). In non-APL AMLs, with 32 ITD and 17 TKD, overall FLT3 mutations were still significantly enriched in obese patients ((27/49 obese vs 22/102 non-obese, Odds Ratio=2, p=0.04) but not when analysed separately (p=0.11 for ITD and 0.44 for TKD).

We then attempted to validate this finding in the APL cohorts, for which data on FLT3ITD (table 5). In the pooled analysis (163 mutated patients / 569 total), OR of having a FLT3 ITD was 1.22 (95% CI 1.05-1.43) *per* each 5 kg/m^2 increase. In the individual cohorts, results were significant in the Italian (30/114 mutated, OR 2.35, 95% CI 1.25-4.42) and USA (14/41 mutated, OR 1.44, 95% CI 0.93-2.24) cohorts, but not in the Spanish (119/414 mutated, OR 1.09, 95% CI 0.89-1.33).

Discussion

In this work we provide substantial evidence for an association between elevated BMI and risk of developing AML. The risk was particularly high with the APL subtype, with an

estimated 44% HR increase *per* each 5 kg/m^2. This was qualitatively confirmed by comparing expected *vs* observed BMI distributions in APL cohorts across three western countries (US, Spain and Italy) with different obesity prevalence and dietary habits. Additionally, we provide hypothesis-generating evidence for molecular mechanisms underlying such association, in particular the possible involvement of pro-inflammatory fatty acid metabolism and mutations of the tyrosin kinase FLT3.

Our epidemiological results expand a growing body of literature identifying overweight/obesity as a bona fide risk factor for leukemias. The most recent meta-analysis reported an adjusted relative risk of 1.14 (95% CI, 1.04–1.26, P = 0.008) for acute myeloid leukemias overall[5]. Despite increasing evidence, the notion of obesity as a risk factor for leukemia remains widely overlooked[7]. Among the highly heterogeneous AML subtypes, APL is the most clinically and biologically coherent. We and others previously showed that in APL, but not in other AMLs, an elevated BMI significantly affects outcome [10,19]. This is also in line with the few retrospective studies that have assessed APL as a separate disease entity [20,21]. No study had addressed this question prospectively, a task made difficult by the rarity of the disease but made possible in our case by the very large study population (5.2 million). The largest prospective study to date (EPIC), which revealed a statistically significant higher risk only in female AMLs, but not in other gender and biological subgroups[4], was based on a relatively small number of incident cases, only 671 out of 375,021 participants over 11.5 years of median follow-up. The use of orthogonal epidemiological approaches is a strength of the study, as it attempts to mitigate some weaknesses of each design. Registry-based studies have little patient selection bias, providing results that are more comparable to real-life scenarios. However, the quality of case identification is likely to be sub-optimal; erroneous assignment of APL to the general AML ICD code might "deplete" incident cases and further reduce statistical power. Case-control studies in the context of clinical trials, on the other

hand, offer the advantage of gold-standard diagnosis but might be affected by significant patient selection biases. This may have counter-selected obese patients in the present study, since the correlated comorbidities may be associated with limited access to clinical trials. Another limitation of the study is that we could not provide the same degree of geographical homogeneity for control subjects in the case-control studies. This may be particularly relevant for the US, known to have ample state-specific differences in BMI distribution. However, this variation is mainly due to demographic parameters[22], such as age, sex and race, and is therefore at least partly accounted for in our multivariate analysis. We also note that our USA APL cohort includes a single patient of hispanic ethnicity. Hispanics are considered at higher incidence of APL, although some large studies based on Surveillance, Epidemiology, and End Results (SEER) data dispute this commonly held conclusion [23]

Understanding the molecular mechanism causing increased cancer risk in obese subjects is crucial for adequate nutritional prevention, given the sustained rise of obesity worldwide, particularly in emerging economies. The possibility to match transcriptional and mutational profiles from TCGA to patient clinical and BMI data provided an opportunity to generate hypotheses grounded on actual data. However, extracting biological significance from large molecular datasets remains challenging. Shifting analytical focus from single genes to gene sets or pathways may allow to capture signals even when the changes affecting individual genes are minimal, provided they are coherent. The gene set-based method we employed here for transcriptional analysis does not assume equal variances, resulting in improved sensitivity and specificity over similar competing methods[15]. Our main finding is the upregulation of several genes involved in the metabolism of pro-inflammatory ω-6 polyunsaturated fatty acids (PUFA, linoleic and arachidonic) in APL. These molecules are increased in the plasma of metabolically impaired subjects, including the obese[24], and may lead to elevated production of derivative molecules with multiple effects in signaling and inflammation, enhancing

leukemogenesis through several independent mechanisms: direct growth promotion, generation of genotoxic oxidative stress, immune modulation[18,25] and generation of endogenous agonists for Peroxisome proliferator-activated receptors (PPAR)[26]. PPARs are known insulin sensitizers [27] and their transcriptional targets are upregulated in APL (Supplementary Table S2); APL expressed higher levels of insulin and IGF1 receptors and its growth may thus be favoured by the increased insulin/IGF1 levels in obese subjects[3,28]. Elevated generation of PUFA-derived eicosanoids by APL cells may also explain the association between obesity and ATRA differentiation syndrome (DS) [10], as eicosanoids strongly promote leukocyte adhesion and chemokine release in the lungs [29].

Finally, the association between FLT3 mutations and a higher BMI, although unconfirmed in the larger Spanish cohort, is an intriguing finding that we think deserves additional research. FLT3 mutations are associated with specific metabolic dependencies which may be differentially affected by the systemic nutritional state[30]. It cannot be entirely ruled out that geographical differences in dietary composition may account for the discrepancies in the association between BMI and APL risk (weakest in Spain) and FLT3 mutations (null in Spain). Consistent with this highly speculative view, a recent EPIC substudy revealed marked differences in nutritional patterns between European nations. Despite sharing a theoretical propensity for "Mediterranean" diets, Italy and Spain were highly polarized especially in terms of average polyunsaturated fatty acid consumption (3 vs 38% of the participants in the highest quintile) [31].

In conclusion, based on evidence provided here, we propose to include obesity among environmental factors increasing risk for myeloid neoplasms and in particular APL. Additional studies with experimental models will clarify the molecular determinants of this relationship and test whether and how specific nutritional components like PUFAs can

determine specific mutational and transcriptional alterations able to influence teh natural history of the disease.

Declaration of interests

The authors declare no relevant conflict of interest

References

1 Mazzarella L, Riva L, Luzi L, Ronchini C, Pelicci PG. The genomic and epigenomic landscapes of AML. Semin Hematol. 2014;51(4):259–272.

2 Shlush LI, Zandi S, Itzkovitz S, Schuh AC. Aging, clonal hematopoiesis and preleukemia: not just bad luck? Int J Hematol. 2015;102(5):513–522.

3 Lauby-Secretan B, Ph D, Scoccianti C, Ph D, Loomis D, Ph D. Body Fatness and Cancer — Viewpoint of the IARC Working Group. N Engl J Med. 2016;375(8):794–798.

4 Saberi Hosnijeh F, Romieu I, Gallo V, et al. Anthropometric characteristics and risk of lymphoid and myeloid leukemia in the European Prospective Investigation into Cancer and Nutrition (EPIC). Cancer Causes Control. 2013;24(3):427–438.

5 Li S, Chen L, Jin W, et al. Influence of body mass index on incidence and prognosis of acute myeloid leukemia and acute promyelocytic leukemia: A meta-analysis. Sci Rep. 2017;7(1):17998.

6 Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. Lancet. 2014;384(9945):755–765.

7 American Cancer Society. Leukemia–Acute Myeloid (Myelogenous). http://www.cancer.org/cancer/leukemia-acutemyeloidaml/detailedguide/leukemia-

acute-myeloid-myelogenous-risk-factors.

8 Komanduri K V, Levine RL. Diagnosis and Therapy of Acute Myeloid Leukemia in the Era of Molecular Risk Stratification. Annu Rev Med. 2016;67:59–72.

9 Coombs CC, Tavakkoli M, Tallman MS. Acute promyelocytic leukemia: where did we start, where are we now, and the future. Blood Cancer J. 2015;5:e304.

10 Breccia M, Mazzarella L, Bagnardi V, et al. Increased BMI correlates with higher risk of disease relapse and differentiation syndrome in patients with acute promyelocytic leukemia treated with the AIDA protocols. Blood. 2012;119(1):49–54.

11 Ley TJ, Miller C, Ding L, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059–2074.

12 Avvisati G, Lo-Coco F, Paoloni FP, et al. AIDA 0493 protocol for newly diagnosed acute promyelocytic leukemia: very long-term results and role of maintenance. Blood. 2011;117(18):4716–4725.

13 Lo-Coco F, Avvisati G, Vignetti M, et al. Front-line treatment of acute promyelocytic leukemia with AIDA induction followed by risk-adapted consolidation for adults younger than 61 years: results of the AIDA-2000 trial of the GIMEMA Group. Blood. 2010;116(17):3171–3179.

14 Adès L, Sanz MA, Chevret S, et al. Treatment of newly diagnosed acute promyelocytic leukemia (APL): a comparison of French-Belgian-Swiss and PETHEMA results. Blood. 2008;111(3):1078–1084.

15 Yaari G, Bolen CR, Thakar J, Kleinstein SH. Quantitative set analysis for gene expression: A method to quantify gene set differential expression including gene-gene correlations. Nucleic Acids Res. 2013;41(18):1–11.

16 Morishima K, Tanabe M, Furumichi M, Kanehisa M, Sato Y. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2016;45(D1):D353–

D361.

17 Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, et al. IntOGen-mutations identifies cancer drivers across tumor types. Nat Methods. 2013;10(11):1081-1082.

18 Wang D, Dubois RN. Eicosanoids and cancer. Nat Rev Cancer. 2010;10(3):181–193.

19 Castillo JJ, Mulkey F, Geyer S, et al. Relationship between obesity and clinical outcome in adults with acute myeloid leukemia: A pooled analysis from four CALGB (alliance) clinical trials. Am J Hematol. 2016;91(2):199–204.

20 Wong O, Harris F, Yiying W, Hua F. A hospital-based case-control study of acute myeloid leukemia in Shanghai: Analysis of personal characteristics, lifestyle and environmental risk factors by subtypes of the WHO classification. Regul Toxicol Pharmacol. 2009;55(3):340–352.

21 Estey E, Thall P, Kantarjian H, Pierce S, Kornblau S, Keating M. Association between increased body mass index and a diagnosis of acute promyelocytic leukemia in patients with acute myeloid leukemia. Leukemia. 1997;11(10):1661–1664.

22 Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. JAMA. 2012;307(5):491–497.

23 Matasar MJ, Ritchie EK, Consedine N, Magai C, Neugut AI. Incidence rates of acute promyelocytic leukemia among Hispanics, blacks, Asians, and non-Hispanic whites in the United States. Eur J Cancer Prev. 2006;15(4):367–370.

24 Caspar-Bauguil S, Fioroni A, Galinier A, et al. Pro-inflammatory phospholipid arachidonic acid/eicosapentaenoic acid ratio of dysmetabolic severely obese women. Obes Surg. 2012;22(6):935–944.

25 Trabanelli S, Chevalier MF, Martinez-Usatorre A, et al. Tumour-derived PGD2 and NKp30-B7H6 engagement drives an immunosuppressive ILC2-MDSC axis. Nat

Commun. 2017;8(1):593.

26 Kliewer SA, Lenhard JM, Willson TM, Patel I, Morris DC, Lehmann JM. A prostaglandin J2 metabolite binds peroxisome proliferator-activated receptor γ and promotes adipocyte differentiation. Cell. 1995;83(5):813–819.

27 Peters JM, Shah YM, Gonzalez FJ. The role of peroxisome proliferator-activated receptors in carcinogenesis and chemoprevention. Nat Rev Cancer. 2012;12(3):181–195.

28 Poloz Y, Stambolic V. Obesity and cancer, a case for insulin signaling. Cell Death Dis. 2015;6:e2037.

29 Dahlén SE, Björk J, Hedqvist P, et al. Leukotrienes promote plasma leakage and leukocyte adhesion in postcapillary venules: in vivo effects with relevance to the acute inflammatory response. Proc Natl Acad Sci U S A. 1981;78(6):3887–3891.

30 Stockard B, Garrett T, Guingab-Cagmat J, Meshinchi S, Lamba J. Distinct Metabolic features differentiating FLT3-ITD AML from FLT3-WT childhood Acute Myeloid Leukemia. Sci Rep. 2018;8(1):5534.

31 Moskal A, Pisa PT, Ferrari P, et al. Nutrient Patterns and Their Food Sources in an International Study Setting: Report from the EPIC Study. PLoS One. 2014;9(6):e98647.

Tables

Table 1. Hazard ratios from the UK population study

Disease	No. of events	Adjusted HR (per 5kg/m^2 increase in BMI)	95% CI
APL	26	1.44	1.00, 2.08
AML	1012	1.17	1.10, 1.26
LL	2823	1.04	1.00, 1.09
Other	1972	1.10	1.04, 1.15

Table 2. **Description of the cross-sectional cohorts**

		Italy	Spain	USA
		No. 134	No. 414	No. 42
Age	18 - 35	46 (34.3)	113 (27.3)	13 (31.0)
	36 - 50	34 (25.4)	145 (35.0)	12 (28.6)
	51 - 65	40 (29.9)	102 (24.6)	11 (26.2)
	> 65	14 (10.4)	54 (13.0)	6 (14.3)
	Median (IQR)	45 (31-57)	45 (34-57)	47 (33-60)
Gender	Male	63 (47.0)	227 (54.8)	21 (50.0)
	Female	71 (53.0)	187 (45.2)	21 (50.0)
Year of diagnosis	Median (range)	2002 (1997-2010)	2003 (1996-2012)	2007 (2001-2011)
Race	White	-	-	36 (85.7)
	Black	-	-	5 (11.9)
	Hispanic	-	-	1 (2.4)
BMI	Median (IQR)	26 (23-28)	26 (23-29)	34 (28-39)

IQR: Interquartile range

Table 3. Observed BMI distribution in APL cases and expected BMI distribution in general population (percentages in brackets)

		All			Males			Females		
Italy	**BMI**	**Obs**	**Exp[a]**	**P-val**	**Obs**	**Exp[c]**	**P-val**	**Obs**	**Expd[c]**	**P-val**
	<25.0	48 (35.8)	77.8 (58.0)	<0.001	16 (25.4)	29.0 (46.0)	<0.001	32 (45.1)	48.8 (68.7)	<0.001
	25.0-29.9	71 (53.0)	44.8 (33.4)		42 (66.7)	28.0 (44.5)		29 (40.8)	16.8 (23.6)	
	>=30.0	15 (11.2)	11.4 (8.5)		5 (7.9)	6.0 (9.5)		10 (14.1)	5.5 (7.7)	
	Total	134	134		63	63		71	71	
Spain	**BMI**	**Obs**	**Exp[a]**	**P-val**	**Obs**	**Exp[c]**	**P-val**	**Obs**	**Expd[c]**	**P-val**
	<25.0	172 (41.5)	189.9 (45.9)	0.011	79 (34.8)	85.0 (37.4)	0.130	93 (49.7)	104.9 (56.1)	0.033
	25.0-29.9	156 (37.7)	158.1 (38.2)		99 (43.6)	103.2 (45.5)		57 (30.5)	55.9 (29.4)	
	>=30.0	86 (20.8)	66.0 (15.9)		49 (21.6)	38.8 (17.1)		37 (19.8)	27.2 (14.6)	
	Total	414	414		227	227		187	187	
USA	**BMI**	**Obs**	**Exp[b]**	**P-val**	**Obs**	**Exp[c]**	**P-val**	**Obs**	**Expd[d]**	**P-val**
	<25.0	2 (4.8)	12.8 (30.6)	<0.001	1 (4.8)	5.3 (25.4)	0.002	1 (4.8)	7.5 (35.7)	0.003
	25.0-29.9	13 (31.0)	13.7 (32.6)		5 (23.8)	7.9 (37.7)		8 (38.1)	5.8 (27.6)	
	30.0-34.9	12 (28.6)	8.6 (20.5)		9 (42.9)	4.8 (23.0)		3 (14.3)	3.8 (18.0)	
	>=35.0	15 (35.7)	6.9 (16.3)		6 (28.6)	2.9 (13.9)		9 (42.9)	3.9 (18.8)	
	Total	42	42		21	21		21	21	

Expected frequencies were obtained from the BMI distribution in the general population of the area of the APL cases, period of APL diagnosis and in addition: [a]age class and sex; [b]age class, sex and race; [c]age class; [d]age class and race.

Table 4. Significantly upregulated KEGG pathways in APL vs AML in TCGA

pathway.name	log.fold.change	p.Value	FDR
KEGG_RENIN_ANGIOTENSIN_SYSTEM	0.6503	0.0023	0.0093
KEGG_LINOLEIC_ACID_METABOLISM	0.6381	0.0002	0.0010
KEGG_GLYCOSAMINOGLYCAN_BIOSYNTHESIS_HEPARAN_SULFATE	0.4217	0.0000	0.0000
KEGG_GLYCOSPHINGOLIPID_BIOSYNTHESIS_LACTO_AND_NEOLACTO_SERIES	0.3391	0.0003	0.0017
KEGG_ALANINE_ASPARTATE_AND_GLUTAMATE_METABOLISM	0.3258	0.0009	0.0039
KEGG_ARACHIDONIC_ACID_METABOLISM	0.3221	0.0037	0.0130
KEGG_GLYCOSAMINOGLYCAN_DEGRADATION	0.3208	0.0000	0.0001
KEGG_HISTIDINE_METABOLISM	0.2996	0.0044	0.0148
KEGG_ARGININE_AND_PROLINE_METABOLISM	0.2582	0.0001	0.0008
KEGG_LIMONENE_AND_PINENE_DEGRADATION	0.1662	0.0087	0.0250
KEGG_CARDIAC_MUSCLE_CONTRACTION	0.1475	0.0084	0.0245
KEGG_PROTEIN_EXPORT	0.1439	0.0066	0.0204
KEGG_PATHWAYS_IN_CANCER	0.1346	0.0169	0.0428

Table 5. Logistic regression of BMI and FLT3 ITD mutations

BMI	**All 3 cohorts 163/569[a]**	**ITALY 30/114[a]**	**SPAIN 119/414[a]**	**USA 14/41[a]**
	OR (95% CI)	OR (95% CI)	OR (95% CI)	OR (95% CI)
5 unit increase	1.22 (1.05-1.43)	2.35 (1.25-4.42)	1.09 (0.89-1.33)	1.44 (0.93-2.24)
≥ 25 *vs* <25[b]	-	4.40 (1.63-11.9)	1.15 (0.75-1.78)	-
≥ 30 *vs* <30[b]	-	-	-	6.46 (1.21-34.5)

[a] Mutations / All patients; [b] Given the small number of obese patients in Italy and Spain, we compared overweight/obese patients *versus* normal weight patients (i.e. BMI ≥ 25 *vs* <25). Given the small number of normal weight patients in USA, we compared obese patients *versus* non-obese patients (i.e. BMI ≥ 30 *vs* <30).

Figure Legends

Fig 1: **relationship between BMI and log-hazard ratio (HR) for leukemias in the UK population.** (**A**) APL; (**B**) other AMLs; (**C**) lymphoid leukemias; (**D**) all leukemias. Mean (dark line) ± 95% confidence intervals (shaded area) is plotted.

Fig 2: **Differential activities of KEGG pathways and insulin/leptin receptors in the M3 vs non M3 quSage comparison in the TCGA.** (**A)** Activity score with 95% CI of 186 KEGG gene sets; significant gene sets are color-coded in red (if upregulated) or green (if downregulated) (**B)** Insulin/IGF1 receptor pathway and leptin receptors. Mean ± 95% confidence interval are plotted

Fig 3. Association between obesity and FLT3 mutations. (A) Bubble plot representing odds ratio vs -logP value of any mutation in 23 driver genes in the TCGA AML cohort. FLT3 (in red) is the only gene with False Discovery Rate < 0.25. Bubble size reflects the number of obese patients with a mutation. Data are tabulated in supplementary table S3

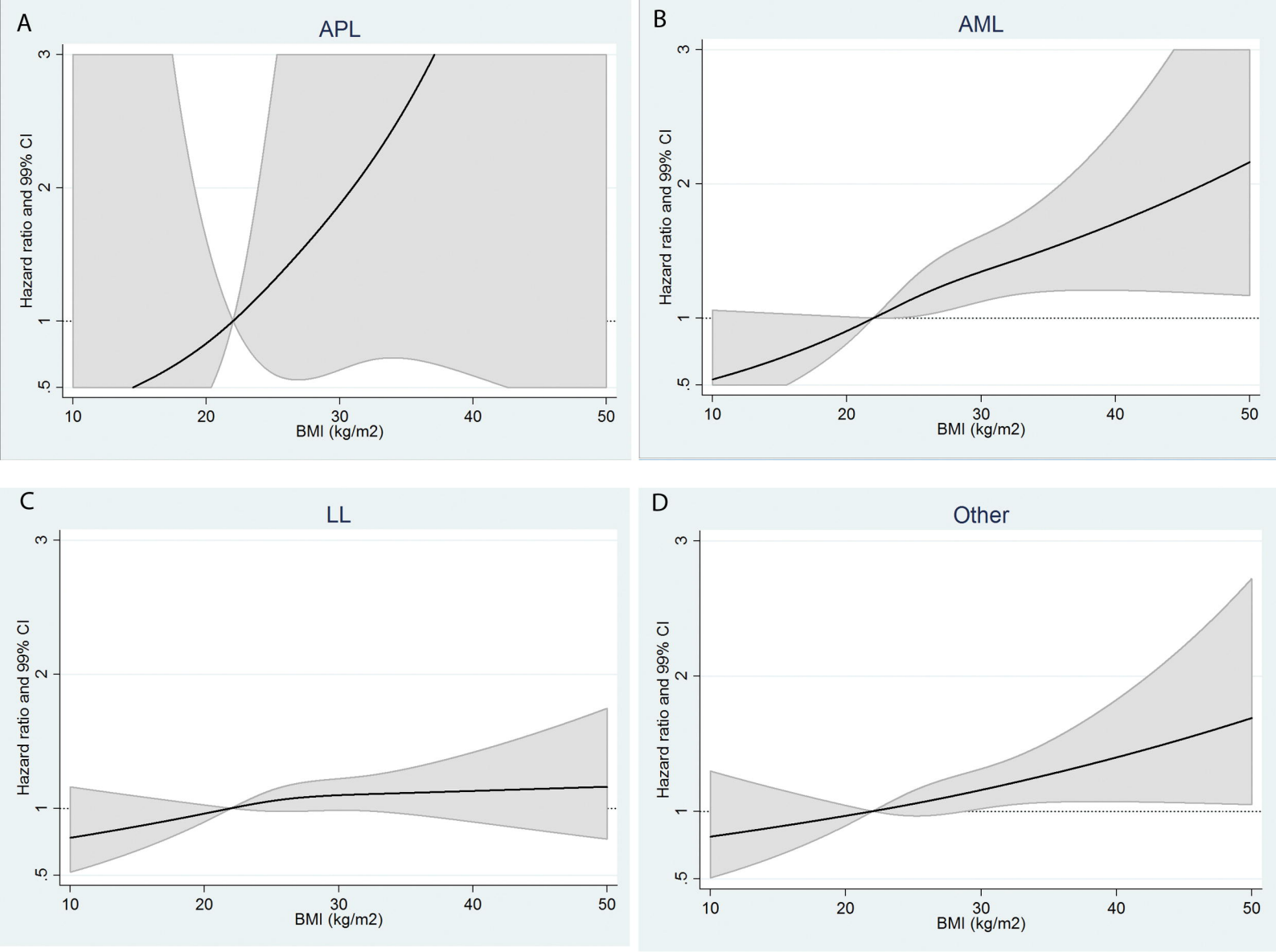
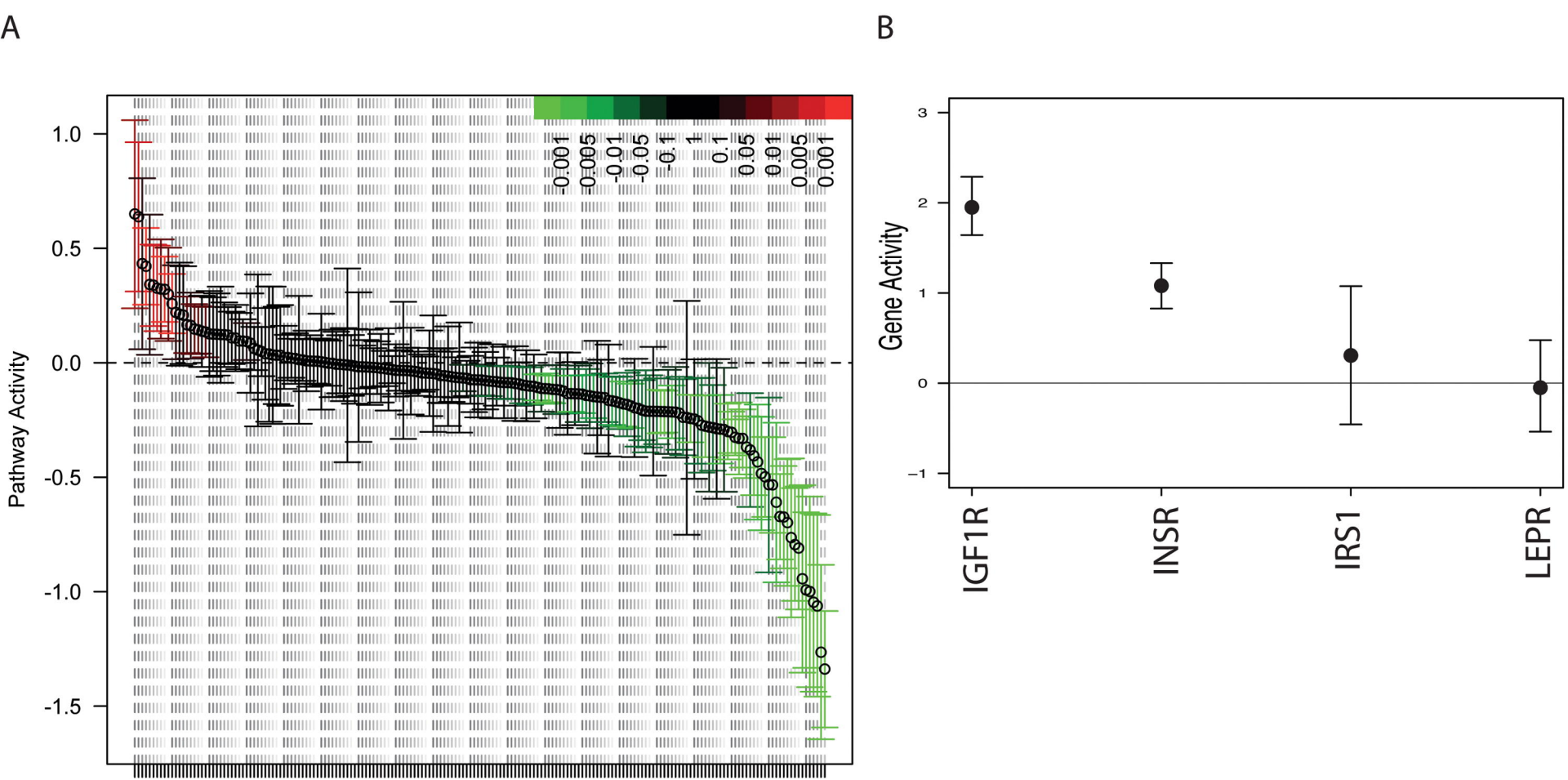

A
APL
Hazard ratio and 99% CI
.5
1
2
3
10
20
30
40
50
BMI (kg/m2)
B
AML
Hazard ratio and 99% CI
.5
1
2
3
10
20
30
40
50
BMI (kg/m2)
C
LL
Hazard ratio and 99% CI
.5
1
2
3
10
20
30
40
50
BMI (kg/m2)
D
Other
Hazard ratio and 99% CI
.5
1
2
3
10
20
30
40
50
BMI (kg/m2)

Figure 1


A
B
Pathway Activity
1.0
0.5
0.0
-0.5
-1.0
-1.5
-0.001
-0.005
-0.01
-0.05
-0.1
1
0.1
0.05
0.01
0.005
0.001
Gene Activity
3
2
1
0
−1
IGF1R
INSR
IRS1
LEPR

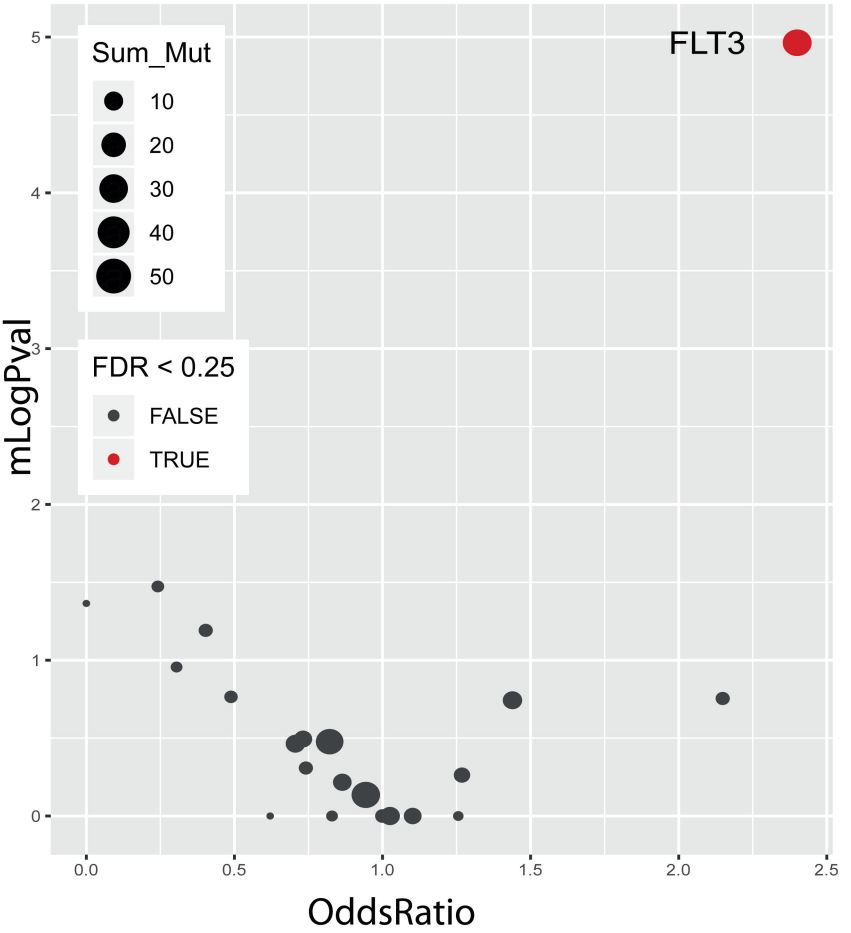

Sum_Mut
10
20
30
40
50
FDR < 0.25
FALSE
TRUE
FLT3
mLogPval
OddsRatio
0
1
2
3
4
5
0.0
0.5
1.0
1.5
2.0
2.5

Fig 3

Supplementary methods

UK population-based study: data collection and statistical methods

Methods for the UK population study were described in depth previously [1]. The study was approved by the London School of Hygiene and Tropical Medicine Ethics Committee. Briefly, data were collected from the UK Clinical Practice Research Datalink (CPRD), which contains computerised primary-care records from general practitioners who use the Vision IT system and have agreed at the practice level to participate (covering about 9% of the UK population). The CPRD dataset is representative of the UK population in terms of age, sex, ethnicity and BMI when compared with census data [2,3] Study entry began 12 months after registration and we assigned BMI records as exposure only 12 months after their recording, to guard against reverse causality (ie, BMI being affected by undiagnosed cancer). We included all people aged 16 years or older with BMI and subsequent eligible follow-up time. BMI was recorded as per local general practice. Individuals with any record of cancer before study entry were excluded. BMI records and diagnosis collected between years 1987-2012 were included in the analysis. To identify outcomes of specific leukaemia sub-types, CPRD clinical records were searched for codes relating to: AML (ICD-10 codes: C92.0, C92.5, C92.6, C93.0, C94.0, C95.0); APL (ICD-10 code C92.4); LL (ICD-10 code C91); and any other leukaemias that were not specifically coded ("other").

Subjects were followed-up from study entry until the earliest of: first cancer diagnoses (any site), death, transfer out of CPRD, or last data collection of the practice. To relate BMI to risk of each type of leukaemia, we fitted Cox regression models with attained age as the underlying

timescale. We fitted fully adjusted models, with BMI as a continuous linear term to estimate the average effect of a 5 kg/m² increase in BMI on leukaemia risk. We also fitted a model including BMI as a 3-knot spline in case of non-linearity in the relationship with leukaemia risk; we tested for evidence of non-linearity by conducting a likelihood ratio test comparing nested models with and without the non-linear terms in the spline basis. We controlled for the following covariates at time of the BMI record(s): age (three-knot restricted cubic spline to allow for non-linearity); smoking status (never smoker, current smoker, ex-smoker); alcohol use (non-drinker, current drinker [light, moderate, heavy, unknown], ex-drinker); previous diabetes diagnosis; index of multiple deprivation (in quintiles, a measure of socioeconomic status); calendar period (<1989, 1990–94, 1995–99, 2000–04, 2005–09, ≥2010); and stratified by sex. We excluded people with missing smoking (49 206/5.24 million [0.9%]) and alcohol status (394 196/5.24 million [7.5%]. All CIs are presented at the 95% level.

Cross-sectional studies: data collection and statistical methods

APL cases from Spain were extracted from the PETHEMA database to include 414 cases diagnosed between 1998 and 2012. APL cases from Italy were 134 adult patients treated with AIDA protocol included in the previously described cohort [4]. APL cases from USA included the entire cohort of the published AML TCGA project [5] (n=20) plus 22 additional APL cases, unselected for any clinical variable, diagnosed at Washington University (Expanded TCGA cohort). For all case cohorts, BMI was measured at the time of diagnosis. Data collection was approved by the Research Ethics Board of each participating institution, as referenced [5–8]

We compared the distribution of BMI observed in the three APL case cohorts to the distribution of BMI expected in the general population of

the same countries. Specifically, to calculate the expected distribution of BMI in Italy we used data from the Italian National Institute of Statistics [9] and we selected the area of Lazio, where the APL cases were diagnosed, in the years 2000-2010. For Spain, we used data from the Eurostat [10] and we selected the general population of Spain in the year 2008, the only year available. For both Italy and Spain, the expected BMI distribution was calculated using the available age- and sex- specific BMI distribution of the general population classified in 3 categories (<25; 25-29.9; ≥30). For USA we used the 2009-2010 data from the American National Health and Nutrition Examination Survey [11]. The expected BMI distribution was calculated using the available race-, age- and sex- specific BMI distribution of the general population classified in 4 categories (<25; 25-29.9; 30.0-34.9; ≥35).

The global null hypothesis that the observed counts did not differ from the expected ones across the BMI categories was tested in a null Poisson regression model, where the observed counts were considered as dependent variable and the expected counts as the offset. We included in the model BMI as an ordinal variable to test the log-linear relationship between BMI and the observed to expected ratio (i.e. to test for linear trend). The Pearson's chi-square goodness of fit test p-value was reported.

Expression data analysis

Expression data (RPKM matrix) were downloaded from the AML TCGA data portal. Cases with available RNAseq, BMI and FAB classification data (177/200) were used in the present study. Cases were classified by FAB in "APL" (FAB="M3") and "non-APL" (FAB ≠ "M3") and by BMI in "obese" (BMI ≥ 30) and "non-obese" (BMI < 30). Genes with < 0.2 RPKM in at least 75% of patients were removed [5]. The Quantitative Set

Analysis for Gene Expression method as implemented in the quSAGE package [12] in the R programming language (v 3.2.3) was used to conduct ***supervised gene set enrichment analysis. For each expressed gene, the quSAGE algorithm calculates a probability density function (PDF) of differential expression between two groups of samples. For each gene set, it then calculates "activity", ie the mean difference in log-expression of individual genes included in a gene set.*** Gene sets with False Discovery Rate (FDR) < 0.05 were considered significant. We focused on the KEGG and CGP gene set collections, downloaded from MSigDB (http://software.broadinstitute.org/gsea/msigdb/). The CGP collection was used to confirm enrichment of previously identified APL-specific gene signatures [13] (supplementary table S2). We focused on the KEGG collection as it is enriched for metabolism-associated gene annotations[14]. The script to generate the present results is available upon request.

Mutational data analysis

For the analysis in the TCGA data, mutational data were retrieved from the TCGA AML paper[5] and AML driver genes were downloaded from IntOgen [15]. For each gene, different mutations were conflated so that gene status in each patient was either "mutated" or "wild type". For each gene we then calculated the number of mutated or wild-type patients in the obese or non-obese groups, and calculated Odds Ratios (OR), 95% confidence intervals (CI) and p-values by Fisher's test with Benjamini-Hochberg correction. Only genes with >1 mutation in the dataset were considered, using the fdsm package in R.

For the analysis of the retrospective cohort, FLT3 Internal Tandem Duplication (ITD) mutational data were provided by the referring centers. Logistic regression was employed to calculate ORs with 95% CI.

Supplementary references

1. Bhaskaran K, Douglas I, Forbes H, et al. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet*. 2014;384(9945):755–765.
2. Herrett E, Gallagher AM, Bhaskaran K, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int. J. Epidemiol.* 2015;44(3):827–836.
3. Bhaskaran K, Forbes HJ, Douglas I, Leon DA, Smeeth L. Representativeness and optimal use of body mass index (BMI) in the UK Clinical Practice Research Datalink (CPRD). *BMJ Open*. 2013;3(9):e003389.
4. Breccia M, Mazzarella L, Bagnardi V, et al. Increased BMI correlates with higher risk of disease relapse and differentiation syndrome in patients with acute promyelocytic leukemia treated with the AIDA protocols. *Blood*. 2012;119(1):49–54.
5. Ley TJ, Miller C, Ding L, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* 2013;368(22):2059–2074.
6. Avvisati G, Lo-Coco F, Paoloni FP, et al. AIDA 0493 protocol for newly diagnosed acute promyelocytic leukemia: very long-term results and role of maintenance. *Blood*. 2011;117(18):4716–4725.

7. Lo-Coco F, Avvisati G, Vignetti M, et al. Front-line treatment of acute promyelocytic leukemia with AIDA induction followed by risk-adapted consolidation for adults younger than 61 years: results of the AIDA-2000 trial of the GIMEMA Group. *Blood*. 2010;116(17):3171–3179.
8. Adès L, Sanz MA, Chevret S, et al. Treatment of newly diagnosed acute promyelocytic leukemia (APL): a comparison of French-Belgian-Swiss and PETHEMA results. *Blood*. 2008;111(3):1078 LP – 1084.
9. ISTAT. Health for all. 2015;
10. European Commission. Eurostat population database. 2016;
11. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999-2010. *JAMA*. 2012;307(5):491–7.
12. Yaari G, Bolen CR, Thakar J, Kleinstein SH. Quantitative set analysis for gene expression: A method to quantify gene set differential expression including gene-gene correlations. *Nucleic Acids Res.* 2013;41(18):1–11.
13. Casorelli I, Tenedini E, Tagliafico E, et al. Identification of a molecular signature for leukemic promyelocytes and their normal counterparts: Focus on DNA repair genes. *Leukemia*. 2006;20(11):1978–88.
14. Morishima K, Tanabe M, Furumichi M, Kanehisa M, Sato Y. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* 2016;45(D1):D353–D361.

15. Gonzalez-Perez A, Perez-Llamas C, Deu-Pons J, et al. IntOGen-mutations identifies cancer drivers across tumor types. *Nat. Methods*. 2013;10:1081.

Table S1. quSage activity scores of KEGG gene sets. Only gene sets with FDR < 0.05 in at least one comparison are shown

	M3vsNonM3.KEGG			ObVSNorm_All.KEGG			ObVSNorm_M3.KEGG			ObVSNorm_nonM3.K EGG		
pathway.name	log.fold.c hange	p.V alue	FD R	log.fold.c hange	p.V alue	FD R	log.fold.c hange	p.V alue	FD R	log.fold.c hange	p.V alue	FD R
KEGG_RENIN_ANGIOTENSIN_SYSTEM	0.6503	0.0 023	0.0 09 3	-0.0187	0.8 880	0.9 85 9	-0.2661	0.5 992	0.6 94 3	-0.0294	0.8 297	0.9 90 1
KEGG_LINOLEIC_ACID_METABOLISM	0.6381	0.0 002	0.0 01 0	0.0537	0.6 317	0.9 85 9	0.1104	0.6 912	0.7 60 7	0.0136	0.9 083	0.9 90 1
KEGG_GLYCOSAMINOGLYCAN_BIOSYN THESIS_HEPARAN_SULFATE	0.4217	0.0 000	0.0 00 0	0.1973	0.0 022	0.4 09 8	0.2753	0.1 104	0.2 23 3	0.1680	0.0 122	0.9 90 1
KEGG_GLYCOSPHINGOLIPID_BIOSYNT HESIS_LACTO_AND_NEOLACTO_SERIES	0.3391	0.0 003	0.0 01 7	0.0490	0.4 612	0.9 85 9	0.0944	0.5 590	0.6 64 1	0.0264	0.7 106	0.9 90 1
KEGG_ALANINE_ASPARTATE_AND_GL UTAMATE_METABOLISM	0.3258	0.0 009	0.0 03 9	0.0366	0.5 529	0.9 85 9	0.2905	0.1 212	0.2 37 2	-0.0065	0.9 168	0.9 90 1
KEGG_ARACHIDONIC_ACID_METABOLI SM	0.3221	0.0 037	0.0 13 0	0.0907	0.2 234	0.9 85 9	0.0785	0.6 402	0.7 26 4	0.0751	0.3 437	0.9 90 1
KEGG_GLYCOSAMINOGLYCAN_DEGRA DATION	0.3208	0.0 000	0.0 00	0.1114	0.0 519	0.9 85	-0.0164	0.9 407	0.9 50	0.1077	0.0 707	0.9 90
			1			9			9			1
KEGG_HISTIDINE_METABOLISM	0.2996	0.0 044	0.0 14 8	0.0633	0.3 112	0.9 85 9	0.1274	0.4 197	0.5 47 5	0.0410	0.5 403	0.9 90 1
KEGG_ARGININE_AND_PROLINE_META BOLISM	0.2582	0.0 001	0.0 00 8	-0.0136	0.7 768	0.9 85 9	0.1283	0.2 240	0.3 65 5	-0.0421	0.4 087	0.9 90 1
KEGG_LIMONENE_AND_PINENE_DEGR ADATION	0.1662	0.0 087	0.0 25 0	0.0088	0.8 790	0.9 85 9	0.0364	0.7 788	0.8 13 8	-0.0030	0.9 581	0.9 90 1
KEGG_CARDIAC_MUSCLE_CONTRACTIO N	0.1475	0.0	0.0	0.0321	0.4	0.9	0.1541	0.0	0.1	0.0120	0.7	0.9

		084	24 5		163	85 9		864	98 5		777	90 1
KEGG_PROTEIN_EXPORT	0.1439	0.0 066	0.0 20 4	-0.0374	0.2 974	0.9 85 9	0.1933	0.0 958	0.2 03 1	-0.0688	0.0 632	0.9 90 1
KEGG_PATHWAYS_IN_CANCER	0.1346	0.0 169	0.0 42 8	0.0002	0.9 991	0.9 99 1	0.2432	0.0 112	0.0 92 2	-0.0316	0.4 585	0.9 90 1
KEGG_UBIQUITIN_MEDIATED_PROTEO LYSIS	-0.0801	0.0 163	0.0 42 5	-0.0111	0.6 109	0.9 85 9	0.1777	0.0 128	0.0 92 2	-0.0258	0.2 492	0.9 90 1
KEGG_PANCREATIC_CANCER	-0.0866	0.0 187	0.0 46 3	-0.0124	0.6 339	0.9 85 9	0.1224	0.0 993	0.2 07 4	-0.0213	0.4 381	0.9 90 1
KEGG_INSULIN_SIGNALING_PATHWAY	-0.0874	0.0 170	0.0 42 8	0.0096	0.7 116	0.9 85 9	0.1847	0.0 012	0.0 40 2	-0.0031	0.9 115	0.9 90 1
KEGG_PYRIMIDINE_METABOLISM	-0.0975	0.0 165	0.0 42 5	0.0062	0.8 245	0.9 85 9	0.0774	0.3 112	0.4 63 1	0.0043	0.8 807	0.9 90 1
KEGG_RNA_DEGRADATION	-0.1018	0.0 146	0.0 39 5	-0.0229	0.4 004	0.9 85 9	0.1292	0.1 580	0.2 80 1	-0.0329	0.2 515	0.9 90 1
KEGG_AMYOTROPHIC_LATERAL_SCLE ROSIS_ALS	-0.1077	0.0 002	0.0 01 3	0.0147	0.5 221	0.9 85 9	0.0734	0.2 217	0.3 65 5	0.0147	0.5 431	0.9 90 1
KEGG_NON_SMALL_CELL_LUNG_CANCE R	-0.1128	0.0 010	0.0 04 2	0.0152	0.5 572	0.9 85 9	0.1381	0.0 261	0.1 03 4	0.0090	0.7 385	0.9 90 1
KEGG_CHRONIC_MYELOID_LEUKEMIA	-0.1140	0.0 001	0.0 00 9	0.0155	0.4 669	0.9 85 9	0.2026	0.0 126	0.0 92 2	0.0030	0.8 915	0.9 90 1
KEGG_COLORECTAL_CANCER	-0.1163	0.0 002	0.0 01 0	0.0126	0.5 758	0.9 85 9	0.2530	0.0 000	0.0 01 0	-0.0052	0.8 215	0.9 90 1
KEGG_PEROXISOME	-0.1242	0.0 125	0.0 34 2	0.0283	0.4 070	0.9 85 9	0.1403	0.1 243	0.2 40 8	0.0241	0.5 087	0.9 90 1
KEGG_APOPTOSIS	-0.1347	0.0 008	0.0 03 8	0.0364	0.1 948	0.9 85 9	0.1970	0.0 187	0.1 03 4	0.0279	0.3 395	0.9 90 1
KEGG_PROTEASOME	-0.1371	0.0	0.0	-0.0334	0.2	0.9	0.1156	0.2	0.3	-0.0412	0.2	0.9

		013	05 3		960	85 9		150	63 5		221	90 1
KEGG_REGULATION_OF_ACTIN_CYTOS KELETON	-0.1426	0.0 154	0.0 40	-0.0296	0.4 645	0.9 85	0.1678	0.0 559	0.1 53	-0.0419	0.3 333	0.9 90
			9			9			3			1
KEGG_PHOSPHATIDYLINOSITOL_SIGN ALING_SYSTEM	-0.1478	0.0 024	0.0 09 4	0.0232	0.4 947	0.9 85 9	0.1730	0.0 077	0.0 92 2	0.0164	0.6 496	0.9 90 1
KEGG_LONG_TERM_POTENTIATION	-0.1525	0.0 053	0.0 17 4	-0.0171	0.6 000	0.9 85 9	0.1396	0.0 108	0.0 92 2	-0.0247	0.4 774	0.9 90 1
KEGG_VALINE_LEUCINE_AND_ISOLEUC INE_DEGRADATION	-0.1664	0.0 021	0.0 08 5	-0.0054	0.8 847	0.9 85 9	0.1395	0.3 371	0.4 86 0	-0.0109	0.7 775	0.9 90 1
KEGG_PENTOSE_PHOSPHATE_PATHW AY	-0.1698	0.0 029	0.0 10 5	-0.0559	0.1 519	0.9 85 9	-0.0887	0.2 304	0.3 72 6	-0.0438	0.2 926	0.9 90 1
KEGG_PATHOGENIC_ESCHERICHIA_CO LI_INFECTION	-0.1747	0.0 028	0.0 10 5	0.0001	0.9 983	0.9 99 1	0.2391	0.0 248	0.1 03 4	-0.0145	0.7 640	0.9 90 1
KEGG_WNT_SIGNALING_PATHWAY	-0.1798	0.0 003	0.0 01 8	-0.0202	0.5 735	0.9 85 9	0.2704	0.0 049	0.0 79 8	-0.0399	0.2 867	0.9 90 1
KEGG_MELANOGENESIS	-0.1822	0.0 101	0.0 28 5	0.0017	0.9 737	0.9 99 1	0.2365	0.0 708	0.1 72 8	-0.0120	0.8 199	0.9 90 1
KEGG_ENDOMETRIAL_CANCER	-0.1884	0.0 002	0.0 01 1	0.0121	0.7 239	0.9 85 9	0.1965	0.0 107	0.0 92 2	0.0039	0.9 124	0.9 90 1
KEGG_MISMATCH_REPAIR	-0.1993	0.0 191	0.0 46 8	-0.0195	0.7 019	0.9 85 9	0.3508	0.0 325	0.1 08 0	-0.0461	0.3 817	0.9 90 1
KEGG_RIBOFLAVIN_METABOLISM	-0.2067	0.0 027	0.0 10 2	-0.0038	0.9 237	0.9 87 4	0.0432	0.6 998	0.7 65 7	0.0027	0.9 527	0.9 90 1
KEGG_DNA_REPLICATION	-0.2114	0.0 206	0.0 49 6	-0.0159	0.7 776	0.9 85 9	0.2973	0.0 941	0.2 03 1	-0.0361	0.5 441	0.9 90 1
KEGG_JAK_STAT_SIGNALING_PATHWA Y				0.0150			0.2723			0.0008		

	-0.2119	0.0084	0.024 5		0.7826	0.985 9		0.0166	0.101 5		0.9909	0.991 8
KEGG_DORSO_VENTRAL_AXIS_FORMA TION	-0.2122	0.0045	0.014 9	0.0181	0.7070	0.985 9	0.2959	0.0227	0.103 4	0.0018	0.9728	0.991 8
KEGG_INOSITOL_PHOSPHATE_METAB OLISM	-0.2128	0.0000	0.002	-0.0010	0.9730	0.999 1	0.1404	0.0113	0.092 2	-0.0037	0.9133	0.990 1
KEGG_EPITHELIAL_CELL_SIGNALING_I N_HELICOBACTER_PYLORI_INFECTION	-0.2133	0.0000	0.001	0.0413	0.2478	0.985 9	0.1706	0.0797	0.185 3	0.0403	0.2762	0.990 1
KEGG_RIG_I_LIKE_RECEPTOR_SIGNALI NG_PATHWAY	-0.2148	0.0006	0.027	0.0316	0.3340	0.985 9	0.1770	0.1163	0.232 5	0.0289	0.3786	0.990 1
KEGG_LONG_TERM_DEPRESSION	-0.2202	0.0068	0.020 7	-0.0218	0.6805	0.985 9	0.1495	0.0884	0.200 5	-0.0272	0.6329	0.990 1
KEGG_BETA_ALANINE_METABOLISM	-0.2380	0.0059	0.019 0	-0.0296	0.5572	0.985 9	0.1582	0.4291	0.550 4	-0.0358	0.4898	0.990 1
KEGG_FRUCTOSE_AND_MANNOSE_ME TABOLISM	-0.2433	0.0000	0.000	-0.0113	0.7895	0.985	0.1312	0.0715	0.172	-0.0124	0.7805	0.990
			1			9			8			1
KEGG_DRUG_METABOLISM_OTHER_EN ZYMES	-0.2535	0.0082	0.024 5	0.0270	0.6671	0.985 9	-0.1247	0.5606	0.664 1	0.0563	0.3928	0.990 1
KEGG_O_GLYCAN_BIOSYNTHESIS	-0.2725	0.0103	0.028 6	0.0280	0.6628	0.985 9	0.4616	0.0169	0.101 5	-0.0006	0.9918	0.991 8
KEGG_CALCIUM_SIGNALING_PATHWA Y	-0.2771	0.0001	0.004	-0.0054	0.9034	0.985 9	0.2483	0.0677	0.170 2	-0.0159	0.7335	0.990 1
KEGG_VALINE_LEUCINE_AND_ISOLEUC INE_BIOSYNTHESIS	-0.2834	0.0044	0.014 8	0.0075	0.8942	0.985 9	0.2955	0.1340	0.251 8	-0.0060	0.9138	0.990 1
KEGG_NOD_LIKE_RECEPTOR_SIGNALIN G_PATHWAY	-0.2899	0.0002	0.010	0.0682	0.1886	0.985 9	0.2244	0.0358	0.112 8	0.0688	0.2078	0.990 1

KEGG_CYTOSOLIC_DNA_SENSING_PAT HWAY	-0.3003	0.0 000	0.0 00 1	0.0227	0.5 065	0.9 85 9	0.1045	0.3 320	0.4 86 0	0.0310	0.3 465	0.9 90 1
KEGG_B_CELL_RECEPTOR_SIGNALING_ PATHWAY	-0.3036	0.0 000	0.0 00 0	0.0131	0.7 092	0.9 85 9	0.2192	0.0 125	0.0 92 2	0.0090	0.7 968	0.9 90 1
KEGG_ERBB_SIGNALING_PATHWAY	-0.3255	0.0 000	0.0 00 0	0.0105	0.7 342	0.9 85 9	0.2496	0.0 011	0.0 40 2	0.0043	0.8 823	0.9 90 1
KEGG_GLYCEROLIPID_METABOLISM	-0.3303	0.0 001	0.0 00 4	-0.0405	0.4 091	0.9 85 9	0.2382	0.0 564	0.1 53 3	-0.0509	0.3 134	0.9 90 1
KEGG_FC_GAMMA_R_MEDIATED_PHAG OCYTOSIS	-0.3306	0.0 000	0.0 00 0	-0.0166	0.7 137	0.9 85 9	0.1495	0.0 525	0.1 50 3	-0.0154	0.7 414	0.9 90 1
KEGG_SPHINGOLIPID_METABOLISM	-0.3680	0.0 000	0.0 00 0	0.0174	0.6 685	0.9 85 9	0.3152	0.0 416	0.1 24 8	0.0076	0.8 462	0.9 90 1
KEGG_RETINOL_METABOLISM	-0.3804	0.0 029	0.0 10 5	0.0302	0.7 095	0.9 85 9	0.0697	0.7 325	0.7 92 1	0.0472	0.5 811	0.9 90 1
KEGG_GAP_JUNCTION	-0.4063	0.0 000	0.0 00 1	-0.0223	0.6 911	0.9 85 9	0.3313	0.0 358	0.1 12 8	-0.0360	0.5 254	0.9 90 1
KEGG_CELL_ADHESION_MOLECULES_C AMS	-0.4328	0.0 008	0.0 03 5	0.0313	0.7 135	0.9 85 9	0.5266	0.0 280	0.1 03 4	0.0053	0.9 546	0.9 90 1
KEGG_NON_HOMOLOGOUS_END_JOINI NG	-0.4837	0.0 005	0.0 02 6	0.0143	0.8 625	0.9 85 9	0.4767	0.0 342	0.1 11 5	-0.0058	0.9 434	0.9 90 1
KEGG_TOLL_LIKE_RECEPTOR_SIGNALI NG_PATHWAY	-0.4982	0.0 000	0.0 00 0	0.0443	0.4 457	0.9 85 9	0.3124	0.0 205	0.1 03 4	0.0448	0.4 386	0.9 90 1
KEGG_OLFACTORY_TRANSDUCTION	-0.5325	0.0 062	0.0 19 7	0.1045	0.4 100	0.9 85 9	0.1150	0.7 770	0.8 13 8	0.1333	0.3 200	0.9 90 1
KEGG_CHEMOKINE_SIGNALING_PATH WAY	-0.5333	0.0 000	0.0 00 0	-0.0104	0.8 722	0.9 85 9	0.1242	0.4 531	0.5 69 4	0.0050	0.9 422	0.9 90 1
KEGG_PENTOSE_AND_GLUCURONATE_ INTERCONVERSIONS	-0.6093	0.0 010	0.0 04	0.0156	0.8 942	0.9 85	-0.1942	0.5 218	0.6 32	0.0700	0.5 586	0.9 90

			2			9			0			1
KEGG_BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS	-0.6717	0.0000	0.000 0	-0.0434	0.5171	0.985 9	0.4095	0.0304	0.106 2	-0.0528	0.4121	0.990 1
KEGG_VIRAL_MYOCARDITIS	-0.6738	0.0000	0.000 0	-0.0105	0.8928	0.985 9	0.3705	0.0201	0.103 4	-0.0122	0.8761	0.990 1
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION	-0.6982	0.0000	0.000 0	0.0580	0.5438	0.985 9	0.4852	0.0123	0.092 2	0.0535	0.5840	0.990 1
KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS	-0.7637	0.0000	0.000 1	0.0663	0.5374	0.985 9	0.1498	0.6639	0.748 4	0.1000	0.3574	0.990 1
KEGG_LEISHMANIA_INFECTION	-0.7943	0.0000	0.000 0	0.0247	0.7869	0.985 9	0.2493	0.1456	0.266 7	0.0455	0.6214	0.990 1
KEGG_PRIMARY_BILE_ACID_BIOSYNTHESIS	-0.8093	0.0000	0.000 0	-0.0140	0.8994	0.985 9	0.0694	0.8325	0.850 8	0.0216	0.8437	0.990 1
KEGG_GRAFT_VERSUS_HOST_DISEASE	-0.9436	0.0000	0.000 1	0.1202	0.3917	0.985 9	0.7427	0.0248	0.103 4	0.1099	0.4481	0.990 1
KEGG_TYPE_I_DIABETES_MELLITUS	-0.9923	0.0000	0.000 0	0.0923	0.4402	0.985 9	0.5909	0.0210	0.103 4	0.0968	0.4174	0.990 1
KEGG_ASTHMA	-0.9993	0.0000	0.000 0	0.0057	0.9727	0.999 1	0.4580	0.2233	0.365 5	0.0146	0.9282	0.990 1
KEGG_ALLOGRAFT_REJECTION	-1.0462	0.0000	0.000 0	0.1173	0.4038	0.985 9	0.6057	0.0569	0.153 3	0.1261	0.3788	0.990 1
KEGG_AUTOIMMUNE_THYROID_DISEASE	-1.0623	0.0000	0.000 0	0.0830	0.5357	0.985 9	0.6382	0.0408	0.124 3	0.0856	0.5274	0.990 1
KEGG_INTESTINAL_IMMUNE_NETWORK_FOR_IGA_PRODUCTION	-1.2645	0.0000	0.000 0	-0.0217	0.8626	0.985 9	0.4684	0.1186	0.234 6	-0.0023	0.9859	0.991 8
KEGG_PANTOTHENATE_AND_COA_BIOSYNTHESIS				-0.0035			0.4184			0.0268		

	-1.3380	0.0 000	0.0 00 0		0.9 670	0.9 99 1		0.0 275	0.1 03 4		0.7 322	0.9 90 1

Table S2. quSage activity scores of previously idntified APL-associated signature and PPARG transcriptional targets

	M3vsNonM3.KEGG			ObVSNorm_All.KEGG			ObVSNorm_M3.KEGG			ObVSNorm_nonM3.KEG G		
pathway.name	log.fold.ch ange	p.Va lue	FDR	log.fold.ch ange	p.Va lue	FDR	log.fold.ch ange	p.Va lue	FDR	log.fold.ch ange	p.Va lue	FDR
CASORELLI_ACUTE_PROMYEL OCYTIC_LEUKEMIA_DN	-0.4517	0.00 00	0.0 000	-0.0506	0.29 15	0.9 995	0.3199	0.01 53	0.0 972	-0.0637	0.16 90	0.8 948
CASORELLI_ACUTE_PROMYEL OCYTIC_LEUKEMIA_UP	0.7557	0.00 00	0.0 000	0.1166	0.01 92	0.9 995	0.1573	0.10 24	0.2 172	0.0723	0.07 55	0.8 948
LI_ADIPOGENESIS_BY_ACTIVA TED_PPARG	1.0578	0.00 00	0.0 000	-0.0871	0.47 97	0.9 995	-0.1286	0.68 90	0.7 617	-0.1412	0.24 97	0.8 968
WANG_CLASSIC_ADIPOGENIC_ TARGETS_OF_PPARG	0.6578	0.00 00	0.0 000	0.0057	0.94 87	0.9 995	-0.1266	0.28 91	0.4 158	-0.0167	0.84 42	0.9 787

Table S3. Mutation prevalences in obese ("OB") and non-obese ("NW") patients in the TCGA dataset

ID	OB_MUT	NW_MU T	OB_WT	NW_WT	OR	CI_inf	CI_sup	fisher	mLogPva l	FDR	Sum_Mut
FLT3	33	22	55	88	2.4000	1.2706	4.5335	0.0070	4.9636	0.1607	55
NPM1	23	30	65	80	0.9436	0.5005	1.7791	0.8732	0.1355	1	53
DNMT3A	20	29	68	81	0.8215	0.4269	1.5808	0.6206	0.4771	1	49
IDH2	9	11	79	99	1.0253	0.4049	2.5967	1.0000	0.0000	1	20
IDH1	10	9	78	101	1.4387	0.5577	3.7119	0.4757	0.7430	1	19
RUNX1	7	12	81	98	0.7058	0.2655	1.8758	0.6287	0.4642	1	19
TET2	7	10	81	100	0.8642	0.3150	2.3712	0.8052	0.2167	1	17
TP53	6	10	82	100	0.7317	0.2552	2.0981	0.6100	0.4943	1	16

NRAS	7	8	81	102	1.1019	0.3835	3.1660	1.0000	0.0000	1	15
CEBPA	6	6	82	104	1.2683	0.3944	4.0783	0.7689	0.2628	1	12
WT1	6	6	82	104	1.2683	0.3944	4.0783	0.7689	0.2628	1	12
PTPN11	4	5	84	105	1.0000	0.2604	3.8410	1.0000	0.0000	1	9
KIT	5	3	83	107	2.1486	0.4991	9.2497	0.4703	0.7544	1	8
KRAS	3	5	85	105	0.7412	0.1722	3.1903	0.7348	0.3081	1	8
U2AF1	2	6	86	104	0.4031	0.0793	2.0483	0.3037	1.1916	1	8
STAG2	2	5	86	105	0.4884	0.0924	2.5799	0.4654	0.7649	1	7
PHF6	1	5	87	105	0.2414	0.0277	2.1051	0.2291	1.4735	1	6
ASXL1	2	3	86	107	0.8295	0.1355	5.0763	1.0000	0.0000	1	5
RAD21	1	4	87	106	0.3046	0.0334	2.7754	0.3843	0.9564	1	5
KDM6A	2	2	86	108	1.2558	0.1733	9.0985	1.0000	0.0000	1	4
DIS3	1	2	87	108	0.6207	0.0554	6.9593	1.0000	0.0000	1	3
EZH2	0	3	88	107	0.0000	0.0000	NA	0.2555	1.3645	1	3
SUZ12	0	3	88	107	0.0000	0.0000	NA	0.2555	1.3645	1	3